第38卷 第4期
2021年7月
干 旱 区 研 究
ARID ZONE RESEARCH
Vol. 38 No. 4
Jul. 2021

基于BP-SVM模型的植被变化模拟研究

贾松涛[1]， 黄生志[1]， 王 浩[2]， 李紫妍[1]， 黄 强[1]， 梁 浩[1]

（1. 西安理工大学，西北旱区生态水利国家重点实验室，陕西 西安 710048；2. 中国水利水电科学研究院，流域水循环模拟与调控国家重点实验室，北京 100038）

摘 要：植被是连接生物圈、大气圈、水圈的重要纽带，对流域生态环境和水热状态变化有重要的影响。当前研究大多聚焦于NDVI与气候因子的相关性分析，少数研究预测NDVI时忽视了考虑滞后性以及增加预报因子对提高模型精确度的影响。基于此，本文比较多元线性回归（MLR）、人工神经网络（ANN）和支持向量机模型（SVM），优选出精度较高的模型。在常规因子（降雨和气温）基础上增加了对植被生长有影响的土壤湿度和日照因子，并考虑了不同因子与NDVI之间的滞时效应。研究结果表明：（1）支持向量机模型（SVM）的拟合能力最强，NDVI预测精度最高，泾河、北洛河流域均方根误差均减少1.8%以上。（2）加入土壤湿度、日照等因子，泾河流域模型预测精度提高，泾河流域均方根误差减少8.8%。（3）考虑滞时后，泾河、北洛河流域均方根误差分别减少15%、11%，NDVI预测精度进一步提高，增加了模型的可靠性。预测结果可对未来制定生态保护策略，指导生态修复具有重要参考意义。

关键词：支持向量机；预测精度；预测模型；滞时效应；植被覆盖度

植被是大气、土壤、流域之间的天然纽带，是陆地生态系统的重要组成部分[1]，亦是气候圈、水圈、岩石圈、生物圈及人类活动相互影响相互作用的统一体现[2]。植被覆盖度作为衡量地表植被状况的一个最重要的指标，其在地表植物蒸腾作用、光合作用、土壤水分蒸发等方面发挥着重要作用[3]。其变化作为土地覆盖变化的主要表现形式之一，间接反映了人类活动对自然生态的影响程度，这一影响对水循环过程极其显著。归一化植被指数（NDVI）被作为重要指标之一，能有效的表征地表植被覆盖度及植物生长趋势，其作为遥感近红外波段与红外波段反射率之差与两者之和的比值参数，对植被覆盖状况的表现具有敏感性的量测，也被公认为是表现植被生长状况的有效参数指标[4]。NDVI在不同地区的差异化表现可以反映这些地区生态系统不同程度的生产力发展状况，不同年的NDVI可反映由于气象要素变化所造成的系统生产力的年际变化[5]。因此，NDVI已经被广泛应用于植被时空分布变化及演变的研究中。植被覆盖指数在气候变化环境下，受到多种因素的影响，呈现出复杂的非线性关系，这增加了对NDVI准确性预测的难度。因此，融合多种因素预报NDVI，对提高NDVI预测的精确度十分重要。

目前，很多学者为了探究归一化植被指数与气候要素之间的关联，对二者进行了大量的相关性分析研究。杜加强等[6]在一致性检验基础上，融合GIMMS NDVI数据与MODIS NDVI数据，将短时间序列拓展到长时间序列，分析了气候变化对植被的影响，结果表明气温主要在春秋季促进植被生长，夏季降水量是调节植被生长的主要因子。Wu等[7]研究发现，温度影响高纬度地区植被覆盖度的大小，是植被生长的主要因素之一。陶帅等[8]利用趋势分析和指数分析等方法，分析了植被覆盖度在年尺度和季尺度上的时间变化及未来变化趋势，并对NDVI时空变化特征及相关因子进行交互作用分析，表明植被覆盖度受海拔、年均温、土壤类型及人口密度多种因子共同影响。李登科等[9]采用相关系数、变化斜率等方法分析了陕北长城沿线风沙区植被变化与降水、气温等气候因子之间的关系，发现不同季节NDVI变化幅度不一，降水量是引起NDVI

收稿日期：2020-12-15； **修订日期：**2021-02-10
基金项目：国家自然科学基金项目（51709221），博士后科学基金面上基金（2018M640155）
作者简介：贾松涛（1996-），女，硕士研究生，主要从事干旱区水文水资源研究. E-mail: jst20156458@163.com
通讯作者：黄生志. E-mail: huangshengzhi7788@126.com.

年际波动的主要因子,人类活动等非气候因素也是引起植被变化的原因。对于干旱半干旱地区的植被,NDVI还与土壤的特性,植被的生根作用等有关[10],这些因素共同决定了植物的生长状况及其覆盖度的变化。王伦澈[11]采用碳通量数据对不同类型植被生产力模型进行标定,与MOSDIS算法比较,分析了影响不同类型植被生长的主要因子,发现总辐射和降水是影响研究区域植被生长季节变化的关键因子。虽然以上研究均表明植被覆盖度受多种因子的共同作用,但目前关于NDVI预测研究中多以降雨和气温作为输入因子。张满囤等[12]基于降水和气温数据,结合了3种优化算法(网络搜索法、遗传算法、粒子群算法)对以支持向量机构建的NDVI短期预测模型参数进行了优选,并利用优选的最佳算法从不同角度构建了单项预测模型且进行线型组合对NDVI进行回归预测,结果表明组合模型可对NDVI进行有效预测。Zhou等[13]将NDVI作为因变量,降水与蒸散发作为自变量,输入多元线性回归模型预测NDVI,结果表明基于降水和蒸散发的多元回归模型可以有效预测不同季节植被的动态变化。然而,植被生长状况的优良程度并不是只受到单一气候因子的影响,除降雨和气温,还和多种气候要素的变化有关,受到多种因子的共同作用。基于此,有必要增加与NDVI密切相关的更多的输入因子,这对预测归一化植被指数,提高模型可靠性有重要意义。

土壤湿度作为重要的陆地生态系统输入因子,是水分和能量交换的重要因素,其对植物作物的生长起到决定性的作用[14]。Tian等[15]研究发现,土壤中的含水量影响植被可获得的储水量,在高纬度地区,温度及日照对植被生长状况的驱动力比降水要强。裴志林等[16]基于探测器模型对年最大植被覆盖度空间分布的影响因子进行定量分析,表明影响植被覆盖度的主导因子是降水与土壤类型和土地利用方式的交互作用。刘静[17]用趋势分析、F检验、皮尔逊相关分析与多元线性回归模型从年尺度和季节尺度分析了不同植被覆盖区域气温、降水、湿度、日照等因素对NDVI的作用机制,发现湿度和日照对植被生长有影响。基于此,可考虑选取模型预测因子时增加土壤湿度和日照数据,同时,由于植被生长状况和气候要素之间并非简单的线性关系,因而构建考虑滞时的可靠的融合多种因素的归一化植被指数预测模型,这对未来制定保护生态策略,指导当地生态修复具有重要参考意义。

本文在对比分析了以降水和气温数据构建的支持向量机(SVM)[18]、人工神经网络(ANN)[19]和多元线性回归模型(MLR)[20]模拟植被变化的基础上,融合土壤湿度、日照等因子的信息,构建了考虑不同因子与植被生长状况之间滞后性的NDVI预测模型,以期较好地提高NDVI预测精度,为该地区的农业生产,旱灾预警,生态发展治理等方面提供依据。

1 研究区概况与方法

1.1 研究区概况

泾河与北洛河是渭河流域的两条主要支流(图1),处于我国西北地区的东部,属于大陆性季风气候,降水主要集中在夏季,冬季雨量偏少[21]。流域内气温季节性变化明显。流域土壤类型主要为黄土,其土壤含水量较少。受光照时间较短,年平均日照时长约5 h。主要植被为落叶松、山杏、山桃、刺槐及草地和农作物等。将渭河流域的两条主要支流泾河和北洛河作为研究区。其中,泾河流域灌区面积占泾河流域面积的0.38%。北洛河农业灌溉面积更大,占北洛河流域面积的0.66%。流域地理位置如图1所示。

1.2 数据来源

采用的数据包括气象数据、土壤数据和NDVI数据,分别来源于中国气象数据网(http://data.cma.cn/)、全球陆面数据同化系统的GLDAS产品(http://disc.sci.gsfc.nasa.gov/hydrology/data-holdings)及地理

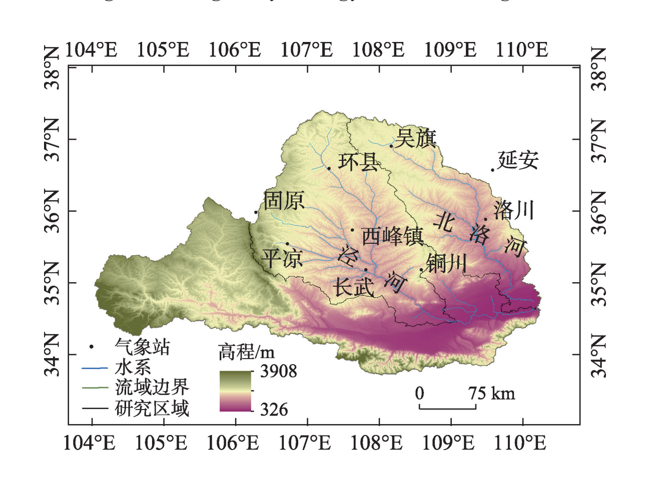


图1 泾河和北洛河流域位置及气象站点空间分布图

Fig. 1 Location and the spatial distribution of meteorological stations of Jing River Basin and Beiluo River Basin

空间数据云（http://www.gscloud.cn）。其中，气象数据为泾河流域环县、固原、平凉、西峰镇、长武和北洛河流域吴旗、延安、洛川、铜川共9个气象站的1982年1月至2013年12月共31年的逐月降水、气温、日照等数据，并利用泰森多边形进行面数据转化。气象站点空间分布如图1所示。

1.3 研究方法

1.3.1 多元线性回归 多元线性回归是用两个或多个自变量解释因变量的一种线性计算模型。多元线性回归模型假设因变量与自变量间存在线性关系[22]，用最小二乘法，计算得到每个自变量对应的回归系数，建立多元线性方程，对自变量实现预报。一般多元线性回归方程如下：

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_mX_m+\varepsilon \quad (1)$$

式中：Y为因变量；β_n 为自变量对应的回归系数，$n=0,1,\cdots,m$；X_n是自变量，$n=0,1,\cdots,m$；ε 则是消除多个自变量对因变量影响的随机误差值。

1.3.2 支持向量机 支持向量机是通过确定核函数，构建最优超平面，解决了复杂的非线性拟合问题。对已知的数据集，支持向量机通过非线性映射，将数据映射到高维特征空间，利用线性问题求解方法在高维特征空间寻求最优的回归函数[23]。其中，线性回归函数如公式(2)：

$$f(x)=(w,x)+b \quad (2)$$

式中：$f(x)$ 为非线性映射函数；w 为权重；x 为第i个样本的输入变量；b 为阈值。

1.3.3 人工神经网络 人工神经网络是一种反向误差传播，与动物神经网类似，不断修正更新权值和阈值的方法，用误差反向传播调整网络参数[24]，是一种被广泛应用具有较强的非线性映射能力和自学习能力的算法[25]。计算公式如下：

$$X_i=f\left(\sum_{k=0}^{n-1}w_{ik}x_k'-\theta_i\right),\quad (i=0,\ 1,\ 2,\ \cdots,\ n-1) \quad (3)$$

$$X_n=f\left(\sum_{i=0}^{m-1}w_{in}X_i-\theta_i'\right),\quad (n=0,\ 1,\ 2,\ \cdots,\ m-1) \quad (4)$$

式中：X_n代表输出层；X_i是隐含层数值；w_{ik}表示从输入层第k个神经元到隐含层第i个神经元权值；x_k'表示输入值；w_{in}表示从隐含层第i个神经元到输出层第n个神经元的权值；θ_i 为隐含层第i个神经元阈值；θ_i' 为输出层第n个神经元阈值。

人工神经网络由输入变量、输出变量、隐含层及激励函数组成。学习过程中，首先对样本进行预处理，确定网络结构，设定权值及阈值。归一化处理数据后，对网络进行训练，模拟网络，得到权值阈值的最优值，计算实际输出值及目标输出间的误差，检验误差，是否达到预设精度，达到则输出结果，未达到则重新模拟网络，计算误差。人工神经网络具有较强针对非线性映射的能力，有较好的自适应学习能力和较强的鲁棒容错等特性[26]。

1.3.4 考虑滞时的多因子预测模型 为了使预测模型更可靠，提高NDVI预测精度，本文考虑了因子的滞后性，并通过添加除降雨、气温以外的其他因子（土壤湿度和日照）的方式，构建了多因子预测模型，以期提高预测模型的精度。

首先，仅将月降雨和月气温数据作为输入变量，分别输入至BP神经网络、支持向量机与多元线性回归模型中比较其预测性能，并优选出最佳预测模型。其次，增加土壤湿度和日照月数据作为输入因子，与只输入降雨及气温的预测结果进行对比，分析预测精度是否有所提高。最后，融合NDVI与降雨、气温、土壤湿度、日照因子的滞时数据并进行预测。

本文选用纳什系数(NSE)[27]、均方根误差(RMSE)[28]和平均相对误差(MRE)[29]构建评价指标体系，评估预测精度。其中，纳什系数验证模型模拟结果的好坏，计算公式如下：

$$\mathrm{NSE}=1-\frac{\sum_{i=1}^{n}(y_i-t_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})} \quad (5)$$

式中：y_i表示观测值；t_i是预测值；$\bar{y}$ 是观测值的平均值。纳什系数趋近于1，表示模型结果好，可信度高，趋于0，则模型结果差，可信度低。

均方根误差表示观测值与真实值之间误差的指标，计算公式如下：

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}(y_i-t_i)^2}{n}} \quad (6)$$

式中：y_i是观测值；t_i是预测值，n为样本数。均方根误差值越小，真实值越接近观测值。

平均相对误差能很好的反应预测值误差的实际情况，计算公式如下：

$$\mathrm{MRE}=\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i-t_i|}{y_i} \quad (7)$$

式中：y_i是观测值；t_i是预测值，n为样本数。流程图如图2所示。

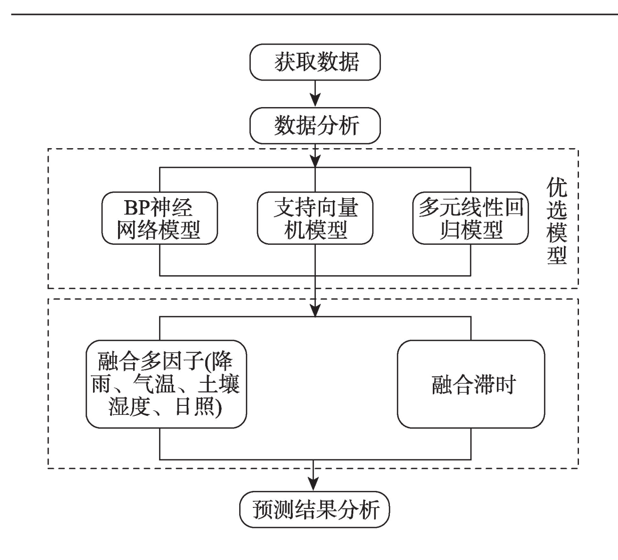


图2 流程图

Fig. 2 Flow diagram

2 结果与分析

2.1 确定最优预测模型

将1982—2013年泾河流域月降雨和气温数据作为输入向量,分别用多元线性回归模型、人工神经网络模型及支持向量机模型对研究区归一化植被指数(NDVI)进行预测,训练期为1982年1月至2006年12月,验证期为2007年1月至2013年12月,输出预测结果,不同模型对NDVI的预测精度各不相同。比较纳什效率系数、均方根误差及平均相对误差,对于不同模型的NDVI预测结果如表1所示。

在泾河流域,用多元线性回归模型预测的结果,验证期其纳什系数仅为0.731,均方根误差为0.092,平均相对误差为0.224;采用人工神经网络模型计算的结果,训练期纳什系数比多元线性回归模型高2.53%;而用支持向量机模型预测的结果,验证期纳什系数最大,为0.761,训练期纳什系数比人工神经网络高1.47%,比多元线性回归模型高4.04%。因此,在泾河流域,支持向量机模型比多元线性回归模型和人工神经网络模型更优,预测精度更高。在北洛河流域,多元线性回归模型预测的结果,验证期纳什系数为0.544,均方根误差0.128,用人工神经网络模型计算,验证期纳什系数为0.611,均方根误差0.120。支持向量机模型预测结果比多元线性回归模型预测结果高21.9%,比人工神经网络模型预测结果高8.52%。训练期支持向量机模型预测NDVI纳什系数与多元线性模型相比,高14.64%,支持向量机模型预测结果纳什系数比人工神经网络模型高0.01%。总体上支持向量机模型仿真模拟结果最好。

结果表明,3种模型中,在处理复杂的非线性映射关系时,人工神经网络模型可利用已有的数据资料及经验,通过复杂的表达方式,表现出输入变量及输出变量之间的非线性关系。人工神经网络模型比多元线性回归模型的预测能力更好,有更优的泛化能力。而在同一环境下,输入相同因子时,支持向量机的优势在于把优化问题转化为凸二次规划问题,而人工神经网络模型在得全局最优解时会出现局部极值的问题,支持向量机模型避免了这一缺点,且相较于多元线性回归模型,其还可以解决复杂的非线性问题。故支持向量机模型的预测精度比前两者更高。

2.2 预测模型输入数据的选择

大多研究只考虑降雨或气温单因子对NDVI的影响,但除降雨和气温,NDVI还和多种气候要素的变化有关,受到多种因子的共同作用。如图3所示,NDVI与不同气候因子和土壤湿度之间均存在复杂的非线性关系。故为提高模型模拟预测结果的准确性,有必要增加与NDVI密切相关的更多的输入因子。本研究将土壤湿度和日照加入降雨气温模型的输入变量中,预测NDVI,将其预测结果与降雨气温作为输入变量的同一模型的预测结果进行比较,由前述研究发现支持向量机模型预测精度更高,故此选用支持向量机作为预测模型。预测结果

表1 多元线性回归、人工神经网络及支持向量机对NDVI预测

Tab. 1 Multiple linear regression, artificial neural network and support vector machine predict NDVI

研究区域	类别	多元线性回归			人工神经网络			支持向量机		
		纳什系数	均方根误差	平均相对误差	纳什系数	均方根误差	平均相对误差	纳什系数	均方根误差	平均相对误差
泾河	训练期	0.820	0.066	0.203	0.936	0.038	0.118	0.928	0.040	0.115
	验证期	0.731	0.092	0.224	0.750	0.088	0.183	0.761	0.086	0.177
北洛河	训练期	0.832	0.077	0.187	0.953	0.040	0.092	0.953	0.040	0.083
	验证期	0.544	0.128	0.205	0.611	0.120	0.191	0.663	0.103	0.184

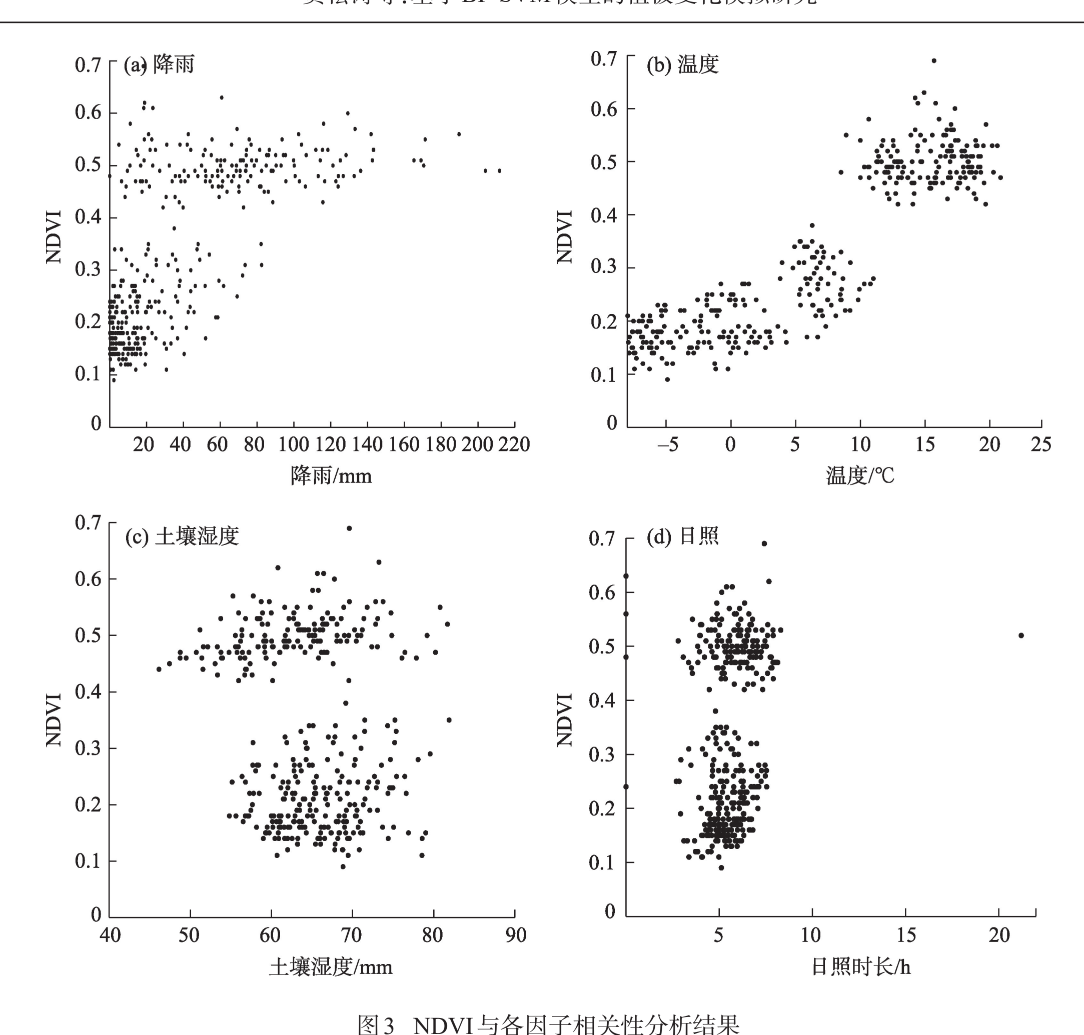


图3 NDVI与各因子相关性分析结果

Fig. 3 Results of correlation analysis between NDVI and various factors

如表2所示。

由表2可知,在泾河,输入变量增加土壤湿度和日照因子后,预测精度提高。只输入降雨和气温两因子预测模型的纳什系数为0.761,而多因子输入变量模型预测结果,其均方根误差比两因子预测模型减少了8.806%,平均相对误差减少5.649%,总体上多因子预测模型精度比两因子预测模型精度要高。在北洛河流域,降雨气温两因子预测模型纳什系数为0.663,多因子预测模型纳什系数比两因子预测模型纳什系数高1.06%,降雨气温两因子预测模型精度与多因子预测模型精度相差较小。

2.3 区域植被对气候因子的滞后性

Wu等[7]利用神经网络算法,构建了NDVI-降雨模型,其结果验证了NDVI与降雨之间存在滞后性,与生态规律一致。故植被生长状况并非只取决于现在的气候条件,与前期气候条件状况有很大的关系,所以本文计算了各个气候因子、土壤湿度与NDVI之间的相关系数。

由不同区域各因子与NDVI之间的相关系数可知(表3),植被变化对气温、降水、土壤湿度及日照

表2 多因子输入模型预测结果比较

Tab. 2 Comparison of prediction results in multiple factors input-model

研究区域	模型类别	降雨气温两因子			降雨、气温、土壤湿度、日照多因子		
		纳什系数	均方根误差	平均相对误差	纳什系数	均方根误差	平均相对误差
泾河	训练期	0.928	0.040	0.115	0.942	0.036	0.099
	验证期	0.761	0.086	0.177	0.729	0.079	0.167
北洛河	训练期	0.953	0.040	0.083	0.959	0.041	0.084
	验证期	0.663	0.103	0.184	0.670	0.118	0.188

表3 不同研究区域各气候因子与NDVI之间相关系数

Tab. 3 Correlation coefficient between various climatic factors and NDVI in different regions

研究区域	气候因子	滞后时间/月												
		0	1	2	3	4	5	6	7	8	9	10	11	12
泾河	NDVI	–	0.80	0.42	0.01	−0.42	−0.78	−0.90	−0.78	−0.41	0.01	0.42	0.78	0.97
	降雨	0.90	0.88	0.62	0.19	−0.30	−0.70	−0.91	−0.88	−0.62	−0.18	0.31	0.69	0.90
	气温	0.68	0.60	0.32	−0.04	−0.35	−0.59	−0.69	−0.61	−0.37	0.01	0.37	0.59	0.68
	土壤湿度	−0.15	−0.32	−0.41	−0.38	−0.23	−0.05	0.15	0.33	0.43	0.41	0.26	0.03	−0.20
	日照	0.20	0.33	0.35	0.31	0.17	0.01	−0.17	−0.30	−0.38	−0.36	−0.24	−0.02	0.18
北洛河	NDVI	–	0.42	0.00	−0.43	−0.79	−0.92	−0.78	−0.42	0.01	0.43	0.79	0.98	0.79
	降雨	0.90	0.87	0.61	0.17	−0.31	−0.71	−0.92	−0.88	−0.61	−0.16	0.32	0.71	0.90
	气温	0.75	0.64	0.33	−0.08	−0.41	−0.64	−0.73	−0.63	−0.38	0.03	0.42	0.66	0.75
	土壤湿度	−0.20	−0.37	−0.45	−0.40	−0.22	−0.01	0.17	0.33	0.42	0.40	0.26	0.01	−0.22
	日照	0.16	0.30	0.32	0.29	0.15	−0.01	−0.15	−0.25	−0.32	−0.31	−0.21	−0.02	0.14

等因子的滞后时间是不同的，这与一些学者的研究结果相符，Anderson等[30]研究了不同气候因子的滞后性，发现植被对气候因子存在响应时间，且对气候因子响应的滞时不同。因此，在探讨气候—植被交互作用机制时，应考虑滞后性。选用气候因子与NDVI之间绝对值大于0.3的相关系数对应的月作为预选滞时，相关系数满足95%置信度检验。因此，泾河流域NDVI滞时是1、2、4、5、6、7、8、10、11、12个月；降雨滞时是0、1、2、5、6、7、8、10、11、12个月；土壤湿度滞时为1、2、3、7、8、9个月；日照滞时选为1、2、3、8、9个月。同理，优选得到北洛河流域NDVI与各气候因子的滞时。

将上述优选结果考虑滞时的NDVI、降雨、气温、土壤湿度、日照各因子作为输入项，以输入因子、目标因子建立支持向量机模型，将其预测结果与未考虑滞后时间影响的预测结果进行比较（表4）。泾河流域，考虑滞时的预测模型其结果，模型的均方根误差比未考虑滞时模型的减小24.39%，平均相对误差减小23.50%；在北洛河流域，考虑滞后时间的模型预测结果纳什系数比不考虑滞后时间的模型结果高34.93%，均方根误差减小51.3%，平均相对误差减小39.46%。对泾河流域和北洛河流域训练期、验证期考虑滞时模型与不考虑滞时模型的预测结果进行比较（图4），考虑各因子、土壤湿度与NDVI之间滞时的模型预测精度比未考虑滞时的模型预测精度显著提高。因此，变化环境下考虑滞后时间的模型对植被生长状况预测是有潜力的。

3 讨论

上述结果表明，考虑滞时的预测模型精度更高。这可能是由于植物本身有一定的耐受力，前期内部有定量的水分储存，在干旱发生时，植被有相应的应对策略，自身应急机制响应，对气候变化的发生存在反应时间，故其不能对气候变化立即做出回应，其生长状况受到前期气候要素变化的影响较大。本研究在以往影响植被生长的常规因子的研究中增加了土壤湿度、日照等敏感因子，结果表明，在泾河流域降雨、气温基础上融合了土壤湿度和日照数据后模型预测精度提高，该结果可能与植被生

表4 支持向量机模型考虑/不考虑滞时的预测结果

Tab. 4 Prediction results based on support vector machine with/without time lag

研究区域	类别	考虑滞时的预测模型			不考虑滞时的预测模型		
		纳什系数	均方根误差	平均相对误差	纳什系数	均方根误差	平均相对误差
泾河	训练期	0.982	0.020	0.048	0.940	0.037	0.103
	验证期	0.880	0.062	0.140	0.787	0.082	0.183
北洛河	训练期	0.989	0.020	0.038	0.947	0.042	0.085
	验证期	0.927	0.056	0.112	0.687	0.115	0.185

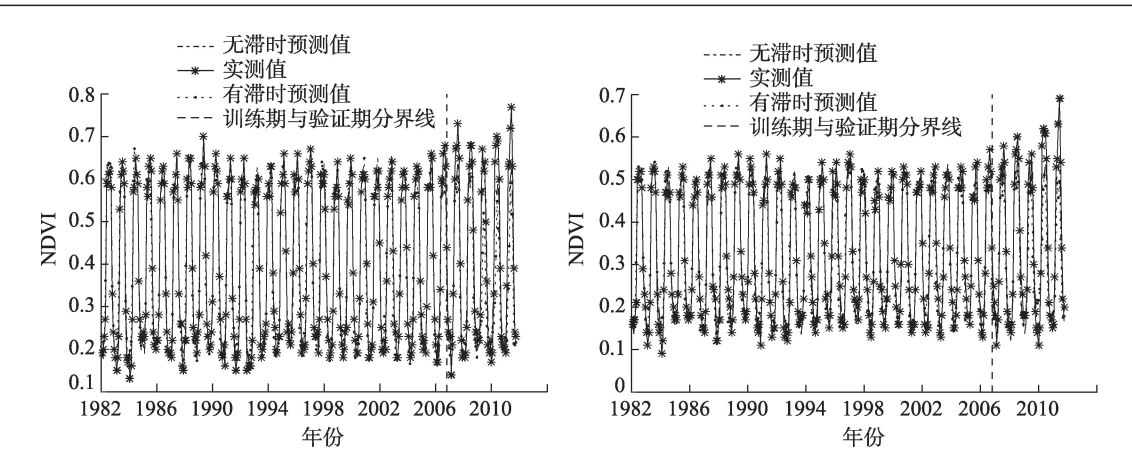


图4 泾河、北洛河流域训练期和验证期NDVI的模拟结果

Fig. 4 Simulation results of NDVI in the training period and verification period of Jing River Basin and Beiluo River Basin

长状况的影响因素并不单一有关：植被覆盖状况在气候变化环境下，受多种因素的影响。而在北洛河，预测精度并无明显变化，可能是由于该研究区域有灌溉面积，泾河研究区灌区面积占流域面积的0.38%。北洛河农业灌溉面积更大，占流域面积的0.66%，受人类活动影响较大。由于不同研究区域NDVI敏感因子不同，所以本研究结果虽不一定适用于所有区域，但本研究运用了多元线性回归、BP神经网络模型和SVM模型3种预测模型，这3种模型对研究区域的选择无要求，各个区域均可适用，所以不同区域的研究，需要优选敏感因子，但3种预测模型和框架可推广到其他区域。通过目前发生的气候变化，为准确预测下一阶段植被生长状况的变化情况，制定应对策略，指导生态修复达到保护生态目的提供理论依据。下一阶段可从3个方面突破：(1) 对NDVI的影响机理深入研究，融合遥相关因子，对输入变量实现预处理。(2) NDVI预测模型的进一步优化，如基于分解技术构建混合模型来提高模型预测精度。(3) 定量分解气候变化与人类活动对NDVI变化的影响。

4 结 论

本文将降雨、气温作为输入变量，分别采用多元线性回归模型、人工神经网络模型和支持向量机模型对渭河流域的支流(泾河、北洛河流域)的NDVI进行了预测，通过分析3种模型的性能优选出适合研究区的最佳预测模型，并融合土壤湿度和日照数据作为输入因子。同时在分析了各输入因子的滞后性对植被生长覆盖程度的影响基础上构建了研究区NDVI考虑滞时的预测模型，得出如下结论：

(1) 3种模型的预测精度由高到低依次为：支持向量机模型>人工神经网络>多元线性回归模型，表明气象要素与植被覆盖度之间存在复杂的非线性关系，且占较大比重。

(2) 支持向量机模型中，输入因子加入土壤湿度和日照数据，精度比降雨、气温两因子预测精度更高，模型更可靠。表明植被生长状况并非受单一因素影响，而受多种因素共同作用。北洛河流域由于受人类活动的影响，精度较无显著变化。

(3) 在对泾河、北洛河研究区植被归一化指数进行模拟预测中，相较于输入无滞时的因子，考虑滞时的预测因子模型预测精度更高。植被受不同气候因素的影响具有不同的滞时，考虑植被主要敏感因子的滞后性对揭示植被生长对气候因素的反应极其重要。

参考文献(References)：

[1] 刘家福, 马帅, 李帅, 等. 1982—2016年东北黑土区植被NDVI动态及其对气候变化的响应[J]. 生态学报, 2018, 38(21): 7647-7657. [Liu Jiafu, Ma Shuai, Li Shuai, et al. Changes in vegetation NDVI from 1982 to 2016 and its responses to climate change in the black-soilarea of Northeast China[J]. Acta Ecologica Sinica, 2018, 38(21): 7647-7657.]

[2] 田甜, 李绍才, 陈敏, 等. 雅砻江流域植被指数长时间序列变化分析[J]. 水力发电学报, 2012, 31(2): 159-164. [Tian Tian, Li Shaocai, Chen Min, et al. Analysis on vegetation index's long time series dynamics of Yalong River basin[J]. Journal of Hydroelectric Engineering, 2012, 31(2): 159-164.]

[3] 杨峰, 李建龙, 钱育蓉, 等. 天山北坡典型退化草地植被覆盖度监测模型构建与评价[J]. 自然资源学报, 2012, 27(8): 1340-1348. [Yang Feng, Li Jianlong, Qian Yurong, et al. Estimating vegetation coverage of typical degraded grassland in the Northern Tianshan Mountains[J]. Journal of Natural Resources, 2012, 27(8): 1340-1348.]

[4] 何辉, 玉素甫江·如素力. 2001—2015年伊犁地区植被NDVI变化及其影响因子的相对作用分析[J]. 中南林业科技大学学报, 2019, 39(10): 76-87. [He Hui, Yusufujiang Rusuli. Analysis of the relative role of vegetation cover changes and its influencing factors in Yili area from 2001 to 2015[J]. Journal of Central South University of Forestry & Technology, 2019, 39(10): 76-87.]

[5] 郭铌, 管晓丹. 植被状况指数的改进及在西北干旱监测中的应用[J]. 地球科学进展, 2007, 22(11): 1160-1168. [Guo Ni, Guan Xiaodan. An improvment of the vegetation condition index with applications to the drought monitoring in Northwest China[J]. Advances in Earth Science, 2007, 22(11): 1160-1168.]

[6] 杜加强, 高云, 贾尔恒• 阿哈提, 等. 近30年新疆植被生长异常值时空变化及驱动因子[J]. 生态学报, 2016, 36(7): 1915-1927. [Du Jiaqiang, Gao Yun, Jiaerheng Ahati, et al. Spatio-temporal patterns and driving factors of vegetation growth anomalies in Xinjiang over the last three decades[J]. Acta Ecologica Sinica, 2016, 36 (7): 1915-1927.]

[7] Wu D H, Zhao X, Huang K C, et al. Time-lag effects of global vegetation responses to climate change[J]. Global Change Biology, 2015, 21(9): 3520-3531.

[8] 陶帅, 邝婷婷, 彭文甫, 等. 2000-2015年长江上游NDVI时空变化及驱动力——以宜宾市为例[J]. 生态学报, 2020, 40(14): 5029-5043. [Tao Shuai, Kuang Tingting, Peng Wenfu, et al. Analyzing the spatio-temporal variation and drivers of NDVI in upper reaches ofthe Yangtze River from 2000 to 2015: A case study of Yibin City[J]. Acta Ecologica Sinica, 2020, 40(14): 5029-5043.]

[9] 李登科, 郭铌, 何慧娟. 陕北长城沿线风沙区植被指数变化及其与气候的关系[J]. 生态学报, 2007, 27(11): 4620-4629. [Li Dengke, Guo Ni, He Huijuan. Vegetation change and its relationship with climate in the region along the Great Wall in northern Shaanxi[J]. Acta Ecologica Sinica, 2007, 27(11): 4620-4629.]

[10] 孙锐, 陈少辉, 苏红波. 黄土高原不同生态类型NDVI时空变化及其对气候变化响应[J]. 地理研究, 2020, 39(5): 1200-1214. [Sun Rui, Chen Shaohui, Su Hongbo. Spatiotemporal variation of NDVI in different ecotypes on the Loess Plateau and its response to climate change[J]. Geographical Research, 2020, 39(5): 1200-1214.]

[11] 王伦澈. 区域大气辐射变化及其对地表植被生产力的定量影响研究[D]. 武汉: 武汉大学, 2015. [Wang Lunche. Reginal Variations of Atmosphere Radiation and its Quantitative Effects on the Terrestrial Ecosystem Productivity[D]. Wuhan: Wuhan University, 2015.]

[12] 张满囤, 黄春萌, 米娜, 等. 基于支持向量机回归的NDVI组合预测模型[J]. 河北工业大学学报, 2017, 46(4): 39-45. [Zhang Mantun, Huang Chunmeng, Mi Na, et al. Combination forecast model of NDVI based on support vector machine regression[J]. Journal of Hebei University of Technology, 2017, 46(4): 39-45.]

[13] Zhou Zhaoqiang, Ding Yibo, Shi Haiyun, et al. Analysis and prediction of vegetation dynamic changes in China: Past, present and future[J]. Ecological Indicators, 2020, 117: 106642.

[14] 杨曦, 武建军, 闫峰, 等. 基于地表温度植被指数特征空间的区域土壤干湿状况[J]. 生态学报, 2009, 29(3): 1205-1216. [Yang Xi, Wu Jianjun, Yan Feng, et al. Assessment of regional soil moisture status based on characteristics of surface temperature vegetation index space[J]. Acta Ecologica Sinica, 2009, 29(3): 1205-1216.]

[15] Tian Siyuan, Van Dijk A I J M, Paul Tregoning, et al. Forecasting dryland vegetation condition months in advance through satellite data assimilation[J]. Nature Communications, 2019, 10: 469.

[16] 裴志林, 杨勤科, 王春梅, 等. 黄河上游植被覆盖度空间分布特征及其影响因素[J]. 干旱区研究, 2019, 36(3): 546-555. [Pei Zhilin, Yang Qinke, Wang Chunmei, et al. Spatial distribution of vegetation coverage and its affecting factors in the upper reaches of the Yellow River[J]. Arid Zone Research, 2019, 36(3): 546-555.]

[17] 刘静. 退耕还林后黄土高原植被覆被变化过程及未来分布预测[D]. 北京: 中国科学院大学, 2019. [Liu Jing. Vegetation Cover Change Process and Future Distribution Prediction on the Loess Plateau after Grain to Green[D]. Beijing: University of Chinese Academy of Science, 2019.]

[18] 陈末, 卢文喜, 侯泽宇, 等. 基于支持向量机的吉林西部地下水水质评价[J]. 节水灌溉, 2013, 38(5): 29-33. [Chen Mo, Lu Wenxi, Hou Zeyu, et al. The assesement of groundewater quality based on support vector machine in Western Jilin[J]. Water Saving Irrigation, 2013, 38(5): 29-33.]

[19] 毛慧慧, 延耀兴, 张杰. 水文预报方法研究现状与展望[J]. 科技情报开发与经济, 2005, 15(19): 172-173. [Mao Huihui, Yan Yaoxing, Zhang Jie. The present situation and prospect of the hydrographic forecasting methods[J]. Journal of Library and Information Science, 2005, 15(19): 172-173.]

[20] 梁浩, 黄生志, 孟二浩, 等. 基于多种混合模型的径流预测研究[J]. 水利学报, 2020, 51(1): 112-125. [Liang Hao, Huang Shengzhi, Meng Erhao, et al. Runoff prediction based on multiple hybrid models[J]. Journal of Hydraulic Engineering, 2020, 51(1): 112-125.]

[21] 代萌, 黄生志, 黄强, 等. 干旱多属性风险动态评估与驱动力分析[J]. 水力发电学报, 2019, 38(8): 15-26. [Dai Meng, Huang Shengzhi, Huang Qiang, et al. Dynamic assessments of drought multi-attribute risks and analysis of its driving force[J]. Journal of Hydroelectric Engineering, 2019, 38(8): 15-26.]

[22] 王秀英, 廖留峰, 王俊杰. 基于多元线性回归的滇西南短时强降水预报模型研究[J]. 气象与环境学报, 2019, 35(2): 15-22. [Wang Xiuying, Liao Liufeng, Wang Junjie. A forecast model for flash heavy rainfall in southwestern Yunnan province based on a multiple linear regression method[J]. Journal of Meteorology and Environment, 2019, 35(2): 15-22.]

[23] 孟二浩, 黄生志, 黄强, 等. 融合大气环流异常因子的径流预报

研究[J]. 水力发电学报, 2017, 36(8): 34−42. [Meng Erhao, Huang Shengzhi, Huang Qiang, et al. Runoff prediction incorporating anomalous atmospheric circulation factors[J]. Journal of Hydroelectric Engineering, 2017, 36(8): 34−42.]

[24] 张俊. 中长期水文预报及调度技术研究与应用[D]. 大连: 大连理工大学, 2009. [Zhang Jun. Mid-and-Long Term Hydrological Forecasting and Operation Techniques Research and Application [D]. Dalian: Dalian University of Technology, 2009.]

[25] 张霞, 李占斌, 张振文, 等. 两种预测模型在地下水动态中的比较与应用[J]. 生态学报, 2012, 32(21): 6788−6794. [Zhang Xia, Li Zhanbin, Zhang Zhenwen. Application and comparison of two prediction models for groundwater dynamics[J]. Acta Ecologica Sinica, 2012, 32(21): 6788−6794.]

[26] 徐冬梅, 赵晓慎. 中长期水文预报方法研究综述[J]. 水利科技与经济, 2010, 16(1): 1−7. [Xu Dongmei, Zhao Xiaoshen. Review on study of mid and long-term hydrological forecasting technique[J]. Water Conservancy Science and Technology and Economy, 2010, 16(1): 1−7.]

[27] 钱剑平. 基于人工神经网络的策勒河流域径流预测研究[D]. 乌鲁木齐: 新疆大学, 2018. [Qian Jianping. A Study on Runoff Prediction in Qira River Basin Based on Artificial Neural Network[D]. Urumqi: Xinjiang University, 2018.]

[28] 杨汉波, 吕华芳, 胡庆芳, 等. 华北平原的大气逆辐射参数化方法比较[J]. 清华大学学报(自然科学版), 2014, 54(5): 590−595. [Yang Hanbo, Lyu Huafang, Hu Qingfang, et al. Comparison of parametrization methods for calculating the downward long-wave radiation over the North China Plain[J]. Journal of Tsinghua University (Science and Technology), 2014, 54(5): 590−595.]

[29] 丛晓红, 拾兵, 于西达, 等. 基于PSO-BP神经网络的黄河利津站输沙量预测[J]. 人民黄河, 2020, 42(1): 1-8. [Cong Xiaohong, Shi Bing, Yu Xida, et al. Prediction of sediment discharge at Lijin station of the Yellow River based on PSO−BP neural network[J]. Yellow River, 2020, 42(1): 1−8.]

[30] Anderson L O, Malhi Y, Aragão Luiz E O C, et al. Remote sensing detection of droughts in Amazonian forest canopies[J]. New Phytologist, 2010, 187(3): 733−750.

Simulation of vegetation change based on BP-SVM mode

JIA Songtao[1], HUANG Shengzhi[1], WANG Hao[2], LI Ziyan[1], HUANG Qiang[1], LIANG Hao[1]

(1. State Key Laboratory of Eco-hydraulics in Northwest Arid Region of China, Xi'an University of Technology, Xi'an 710048, Shaanxi, China; 2. State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China)

Abstract: Vegetation is an important link that connects the biosphere, atmosphere and hydrosphere, and has an important impact on the watershed in the ecological environment and on the exchange of water and heat. Previous studies focused on correlation analyses between the normalized difference vegetation index (NDVI) and climate factors, but the lag and the increase of forecast factors were rarely used to improve the precision of the NDVI model. Thus, this article compared the prediction performance of the multiple linear regression, artificial neural network and support vector machine, then selected the model with the highest precision. Using conventional factors (rainfall and temperature), the prediction performance was tested when soil moisture and sunshine duration were increased, which affect vegetation growth, and the time lag effect between the different factors and NDVI were considered. (1) SVM had the strongest fitting ability and the highest NDVI prediction accuracy. The root-mean-square error of the Jing River Basin and the Beiluo River Basin were reduced by more than 1.8%. (2) The root-mean-square error (RMSE) of Jing River Basin decreased by up to 8.8% when soil moisture, sunshine and other factors were added. (3) Considering the lag time, the root-mean-square error of the Jing River Basin and the Beiluo River Basin decreased by 15% and 11%, respectively, and the predicting accuracy of NDVI was improved further, thus increasing the reliability of the model. These prediction results have an important reference significance for the formulation of ecological protection strategies and the guidance of ecological restoration in the future.

Keywords: support vector machines; prediction accuracy; prediction model; time-lag effect; vegetation coverage